Available online at www.sciencedirect.com
SCIENCE DIRECT
Int. J. Human-Computer Studies 64 (2006) 541–552
International Journal of Human-Computer Studies
www.elsevier.com/locate/ijhcs

Presence, workload and performance effects of synthetic environment design factors

Ruiqi Ma, David B. Kaber*

Edward P. Fitts Department of Industrial Engineering, North Carolina State University, Raleigh, NC 27695-7906, USA

Received 31 July 2005; received in revised form 10 December 2005; accepted 29 December 2005
Available online 2 March 2006
Communicated by K.S. Severinson Eklundh

Abstract

There remains a limited understanding of factors in presence and its relation to performance. This research examined a range of synthetic environment (SE) design features (viewpoint, auditory cue type and visual background) suspected to influence presence, and evaluated differences in presence, workload and task performance caused by manipulations of the factors and task difficulty in a virtual-reality-based basketball free-throw task. Thirty-two research participants were also required to perform secondary-monitoring tasks to assess attention allocation to the virtual and (surrounding) real environments, as an indicator of presence. Analysis of variance results demonstrated immersiveness (viewpoint) and auditory cue type to significantly influence the sense of subjective presence and perceptions of workload. Virtual task performance was significantly affected by task difficultly. This study also provided further evidence of significant positive relations between presence and workload, but no evidence of a correlation of objective presence and performance. These results have general applicability for the design of multimodal SE-based interfaces for real-world tasks, such as telerobot control.

Keywords: Virtual reality; Presence; Multimodal display design; Workload; Secondary-task performance

1. Introduction

Presence has been defined as the sense of being physically present within a computer-generated or remote environment, and has been identified as a design ideal for synthetic environments (SEs) (Draper et al., 1998). When people refer to this subjective experience in the context of using an SE as an interface to remote actuators and sensors to control a distant machine or robot (i.e., perform teleoperation), the terminology ''telepresence'' is more commonly used. Both terms are used in this study; specifically, ''telepresence'' is used in relation to human control of teleoperators and ''presence'' is used in relation to human performance in virtual reality (VR) simulations. In general, an SE is used to provide computer-mediated human interaction with an environment that is physically separate from the user in order to allow human perceptual, cognitive and psycho-motor capabilities to be projected into normally inaccessible, hostile or simulated environments (Draper et al., 1999). Since current technologies do not allow for the development of robots that have the information processing and motor control abilities of a human being, human operators remain involved in operation of robots at remote sites through computers or other advanced technologies for perceptual and cognitive tasks. This reality has caused a change in the direction of teleoperation research from the design of anthropomorphic robots (c. 1970) to a focus on human–computer (machine) interface design that provides high-fidelity displays, including rich visual, auditory and touch information on a remote site in order to facilitate operator perception and a sense of telepresence at the remote site (Riley, 2001). This contemporary direction of teleoperation research was

Abbreviations: ANOVA, analysis of variance; HMD, head-mounted display; ITQ, Immersive Tendencies Questionnaire; PQ, Presence Questionnaire; RW, real world; SD, signal detection; SE, synthetic environment; SSQ, Simulator Sickness Questionnaire; VR, virtual reality

*Corresponding author. Tel.: +1 919 515 3086; fax: +919 515 5281.
E-mail addresses: dbkaber@eos.ncsu.edu, dbkaber@unity.ncsu.edu (D.B. Kaber).

1071-5819/$ - see front matter © 2006 Elsevier Ltd. All rights reserved.
doi:10.1016/j.ijhcs.2005.12.003

based on the belief that the sense of telepresence will improve overall teleoperation (human-in-the-loop) task performance (Sheridan, 1992).

1.1. Factors in presence

With this in mind, it is important to identify and understand factors potentially contributing to the sense of (tele)presence in an SE in order to better describe the phenomenon, predict its occurrence and any relation of the phenomenon to teleoperation or virtual task performance. Many researchers state that a wide range of factors may influence presence from the vividness of an SE or VR simulation (i.e., similarity to the real world (RW)) to the interactivity of an SE (see Nash et al., 2000; Sadowski and Stanney, 2002). For example, Nash et al. (2000) hypothesized that computer system and VR technology, communication mediums, individual differences, and virtual tasks and external environments influence the sense of presence. Their research provides theoretical information that can be used as a basis for designing presence experiments, including specifying independent variables, in order to investigate controlled manipulations of the sense of presence. However, Nash et al. did not provide information on operational methods that could be used to control the sense of telepresence in applications, like teleoperation.

Sheridan (1992) proposed three principal determinants of the sense of presence in an integrated model, including the extent of sensory information, the control of sensors and the ability to modify the computer-generated or remote environment. The sensory information dimension of this model includes visual and auditory channels, viewpoint and other factors, such as tactile feedback, which may be important to providing a sense of telepresence in telerobot operations. Our research was focused on this determinant, as part of Sheridan's model. As an example of prior, related research, a study by Barfield et al. (1995) found that the greater the level of visual realism of an SE, the greater the sense of presence, in general.

Psokta and Davidson (1993) also identified basic requirements for VR system design, including immersion, which is based on the perceptual or visual viewpoint of a user and is thought to facilitate users' perception of VR in a manner similar to their perception of reality. Viewpoint in a VR not only dictates the degree of immersion, but it may also have a significant influence on presence. An egocentric viewpoint, perceived from the perspective of the user, is typically expected to provide a greater sense of self (Slater et al., 1996) in an SE and awareness of objects in the environment, as compared to an exocentric viewpoint, which provides a "third" person perspective.

Beyond visual stimuli, Dinh et al. (1999) also examined the role of other perceptual modalities in presence experiences in VR. They found significant main effects of auditory cues and tactile cues on presence. On the basis of this and the above research on potential factors in presence, it can be contended that VR design should incorporate many sensory cues in order to motivate the sense of presence in any experiments assessing its importance to performance in SEs. There is a need to manipulate this sensory information (e.g., visual cues, viewpoint, auditory cues, etc.) in SEs to examine the sense of (tele)presence and investigate how it relates to virtual (or teleoperation) task performance.

1.2. Measures of presence

In order to evaluate the relations of presence with performance and workload, a valid and objective measure of presence is needed. Since the sense of presence may be determined by many system, task and environmental factors, it is a complex research challenge to develop valid and reliable measures of the phenomenon, and there is likely no single index that will adequately assess the experience of presence (Stanney et al., 1998). Although both subjective (survey-based) and objective (quantitative) measures have been developed and used over the past decade, there are still no universally accepted measures of presence. With respect to subjective measures, Witmer and Singer (1994) developed a Presence Questionnaire (PQ), which includes many items on immersiveness, vividness, completeness and the realism of SEs. This subjective measure has been validated through many empirical studies (e.g., Witmer and Singer, 1998; Riley et al., 2004; Sheik-Nainar et al., 2005). More recently, Sas and O'Hare (2003) developed a presence query technique that incorporated cognitive factors, such as empathy, absorption, creative imagination and willingness to experience presence. Some researchers have also successfully used a few simple questions to measure the sense of telepresence subjectively; e.g., Draper and Blair (1996) used two subjective questions in order to measure telepresence in a teleoperation task and demonstrated significant correlations with mental workload in teleoperation tasks.

Nichols et al. (2000) compared rating scales, such as these, with reflex response and background awareness, as measures of presence in a "duckshoot" virtual environment, and all of the measures were significantly intercorrelated. Slater and Wilbur (1997) also measured presence according to participant physical reactions to startling or unexpected events in a VR. Other objective measures of presence have been proposed by Slater and Steed (2000), including a virtual presence counter, which counted the number of participant mental transitions from VR displays to reality (and vice versa) as a measure of presence. All of these objective measures appeared to be successful in characterizing presence in specific applications. However, they have only been applied in preliminary experiments, and have not been assessed for validated in multiple studies using hypotheses or correlation analyses with other performance measures.

Related to defining objective measures of presence, Draper et al. (1998) presented a structured attentional

resource model of telepresence, defining the concept in terms of concentration on task-relevant and distracter information across local and remote (or real and virtual) environments in a teleoperation scenario. Increases in the allocation of attentional resources to remote (or VR) task information were hypothesized to cause increases in telepresence. That is, a higher attention allocation (or signal detection (SD) rate) in a test VR, compared to the RW, was considered to indicate a higher sense of telepresence. This attention-based formulation of presence has also been investigated as an objective approach to measuring the construct. Riley and Kaber (2001) conducted a study to investigate the relationship between telepresence and attention to VR displays. They found a significant correlation between visual attention to VR displays, versus elements of the surrounding reality, and ratings of telepresence; thus, supporting the use of attention allocation as potential objective indicator of telepresence.

1.3. Presence and task performance

One of the most interesting and challenging tasks in telepresence research is investigation of the relationship between telepresence and task performance. Based on previous teleoperation research, it is critical to establish states of telepresence that may enhance teleoperation system performance (or performance in VRs). It is also important to establish the relation between telepresence and performance in order to develop a model of telepresence for prescribing effective teleoperator designs that may ultimately improve remote task performance (Draper et al., 1998). Studies by Kaber and Riley (2000) and Kaber et al. (2000), involving experiments on simulated telerobot control operations, revealed significant positive relationships between subjective measures of telepresence and performance. In studies like this, performance has usually been measured objectively and telepresence has been measured subjectively using rating techniques. Unfortunately, there have been no correlation studies of objective telepresence and objective measurements of performance. Although the results of previous experiments show a positive correlation between subjective and objective measures of the constructs of telepresence and performance, there is no direct evidence of a causal relationship. The question of whether the sense of presence in VR is related to task performance essentially remains unanswered (Welch, 1999).

1.4. Objectives

The objective of this work was to investigate the impact of the sensory elements in Sheridan's (1992) model on experiences of presence and the potential relation to performance in a virtual task. We examined a range of SE design features, including viewpoint, auditory cue type and visual information, which were suspected to influence presence, and evaluated differences in objective and subjective presence, workload and task performance caused by manipulations of these factors and SE task difficulty. This research also involved correlation studies of objective presence (an attention-based measure) and objective measurements of performance. It was generally hypothesized that as the number and richness of sensory cues in an SE was increased that presence ratings and user allocation of attention to VR displays would increase. It was also hypothesized that an attentional, or SD-based, measure of presence would be sensitive to sensory cue manipulations in the VR and that the measure would correlate with subjective presence ratings and virtual task performance.

2. Method

2.1. Participants

Thirty-two college students from North Carolina State University (NCSU) were recruited for the study and were provided with compensation for their participation. All participants were required to have 20/20 or corrected to normal visual acuity. The average age of the participants was 23.7 years, with a range from 19 to 30 years. As part of an anthropometric data survey, participants were asked to rate their prior experience with VR applications, in playing video games, and using a PC. They were also asked about any experience playing basketball. With respect to VR experience, the average response (on a scale from 1 = "none" to 5 = "frequent") was low (1.7). With respect to playing video games, on average participants indicated moderate experience (3.2). With respect to PC experience, the average participant rating indicated a high frequency of use (4.7). Finally, in regard to playing basketball, on average participants reported relatively infrequent experiences (2.7).

2.2. Experiment task and equipment

The task used in this study was a medium-fidelity, 3-D simulation of a basketball free-throw presented in VR. The simulation presented participants with realistic visuals and sounds of a basketball game environment, a self-representation (i.e., a virtual player) and a virtual basketball. The participants were required to stand and move in a manner similar to shooting a basketball while immersed in the VR; however, they did not handle an actual basketball. The task was expected to allow for a sensitive evaluation of a potential relation between presence and performance because most people are familiar with basketball. The potential for presence experiences in the VR was expected to be greater in the basketball simulation than in, for example, a unique remote robot (teleoperator) control scenario that many participants may not be familiar with. We thought that if a relationship between presence and performance measures could not be established in a virtual

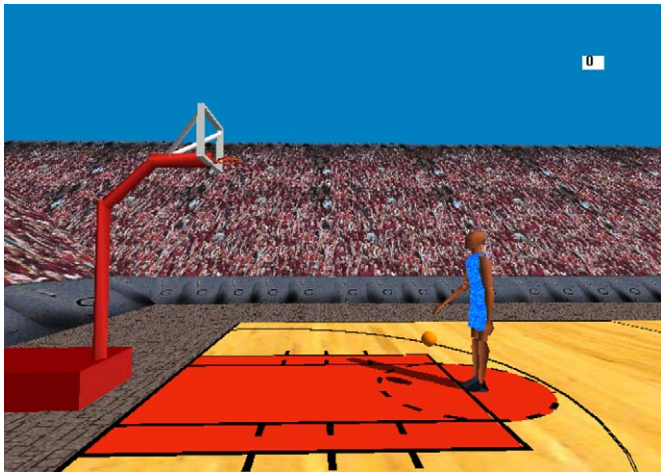

Fig. 1. Basketball simulation with exocentric viewpoint of task (from sideline).

Fig. 2. User immersed in virtual basketball task.

simulation of such a common sporting task, then it may be less likely that a relation would be observed in a complex SE for a teleoperation task. If there was evidence of an actual linkage between presence and performance in the basketball free-throw simulation, then this would motivate additional applied research to investigate the role of multimodal SE design factors in teleoperation scenarios.

The basketball simulation presented a stadium in which a player makes free-throw and 3-point shots (see Fig. 1). The court had all of the conventional lines painted on it, and the goal had a backboard with a rim. There were two different backgrounds available in the simulation. One included gray walls, and the other showed a stadium with a crowd watching the player. The stadium also had a scoreboard. The proportions of object sizes in the simulation were representative of the proportions that would be expected in a real basketball game. (The size of the court was scaled realistically based on measurements of the basketball coliseum at NCSU.)

Different models of an anthropometrically correct basketball player were presented as part of the simulation to ensure compatibility with the handedness of research participants. The models and the virtual stadium were presented to users through a head-mounted display (HMD) integrated with a Silicon Graphics workstation. User control of virtual player (avatar) behaviors occurred via inputs from an Ascension Technologies six degree-of-freedom mouse. The HMD was used to isolate participants' vision to the VR and to simulate 3-D viewing of the virtual environment. However, the blinders, as part of the HMD (adjacent to the sides and bottoms of user's eyes), were taped-up to the body of the HMD so that participants could see the floor of the real lab environment and their feet through the bottom of the helmet.

Participants were asked to shoot virtual basketballs with the six degree-of-freedom mouse and to attempt to make as many baskets/goals as possible (see Fig. 2) in 2-min periods. Participant motions directed the virtual player in shooting the ball. The Ascension Technologies mouse recorded participant hand movements and translated them to movement of the virtual ball. The point of release of the basketball was defined by the computer system at the outset of the experiment. Based on the anthropometry of each participant, when the mouse was held in the hand, the upper arm was held in a horizontal position and the lower arm was positioned vertically (perpendicular to the upper arm), the basketball was released by the virtual player. The faster a participant moved his/her arm, the greater the force applied to the virtual basketball in sending it to the goal. Initially, the virtual basketball player stood in one position. (The position changed later according to the experiment design.) Participants stood on a fixed physical point in the research lab that was close to the simulation workstation. They were not permitted to change their physical position during the experiment. After the experimenter calibrated the arm position of the participants, they repeatedly attempted a specific shot in order to score as many goals as possible. Participants could not control the dribble of the basketball during the experiment.

In addition to shooting basketballs, the participants were required to attend to two secondary tasks during test trials, which involved the detection of random visual cues both in the simulation environment and outside the VR (in the real laboratory). The random cues were modeled as photo-flashes in the stands of the virtual basketball stadium and real strobe light flashes in the research laboratory in which the experiment was conducted. Participants were required to say "flash" when they saw a virtual camera flash in the virtual stadium or "light" when they detected an actual strobe light flash in our lab. The lighting conditions were controlled in the lab with an ambient level of approximately 70 lx. The peak intensity of the strobe light was substantially greater than this allowing for easy detection when attended to. An experimenter recorded participant responses and they were used to calculate the ratio of attention allocation to the VR and RW in the SD tasks.

Sequences of the photoflashes and strobe light flashes were controlled using a software algorithm and programmable logic controller, respectively. They were randomized for each participant and the same sequences were used across participant groups corresponding to virtual environment viewpoint settings (to be discussed). The experimenters knew the timing of all secondary-task cues in advance of trials and they recorded whether participants responded to them or not, as trial time elapsed. The secondary tasks were expected to provide information on the distribution of participant attentional resources across the VR and RW, and it was used as the indicator of presence. Our expectation was that the greater the extent to which participants were immersed in the VR, attending to stimuli from the simulation, and experiencing presence in the basketball task, the less likely they would be to detect visual stimuli in the surrounding reality (the strobe light flashes). In this way, the attention measure was expected to indicate presence.

2.3. Variables

2.3.1. Independent

Four variables were manipulated in the experiment including the VR viewpoint, auditory cues (sounds), visual background and shot distance (task difficulty). These variables were selected for assessing presence and performance relations in the virtual task based on the research by Psokta and Davidson (1993), Barfield et al. (1995), Slater et al. (1996), Dinh et al. (1999) and Kaber and Riley's (2000) demonstration of task difficulty effects on subjective telepresence in a teleoperation simulation. There were four viewpoint conditions including an egocentric view (see Fig. 3 for the virtual player's perspective), an exocentric view from behind the player, an exocentric view from the sideline of the court (see Fig. 1), and a selectable viewpoint. (*Note*: Fig. 1 shows a "rich" visual background that we will describe later. Fig. 3 presents a "simple" visual background.) The selectable viewpoint condition allowed participants to choose from all of the above three viewpoints at their discretion. When using the egocentric viewpoint, participants were able to see virtual representations of the virtual player's arms when making a shot. In the upper right corner of any of these simulation views, a small text box presented the number of baskets/goals made during each experimental trial.

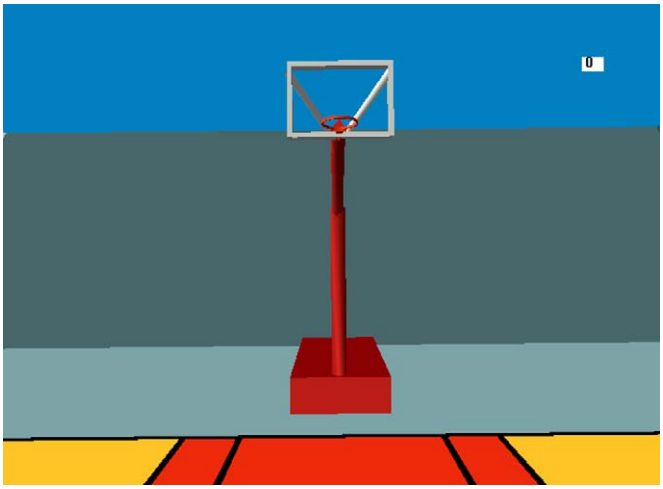

Fig. 3. Egocentric view of simulation with low-fidelity background.

There were also four levels of sound including task-relevant sounds, task-irrelevant sounds, a combination of task-relevant and irrelevant sounds, and no sound. The relevant sounds included the sounds of the ball bouncing on the floor, bouncing off of the backboard or off of the rim, and cheering. There was also a sound to indicate the success of a shot ("swish"). The irrelevant sounds included sounds of crowd noise, sounds of a shot clock and arena music.

There were two visual background conditions including a simple visual of the stadium, as shown in Fig. 3, and a rich visual background stadium, as shown in Fig. 1. The former background was composed of rendered, plain gray, walls surrounding the basketball court, a rendered court and goal post. The rich visual background displayed a texture of a crowd watching the game, textured panels around the court and textured flooring. All textures were produced from photographs of the coliseum at NCSU.

Each participant experienced two levels of task difficulty, including 2-point and 3-point shot distances, during the course of the experiment. The distance varied randomly for each sound and background condition. There were 2-point and 3-point avatar positions programmed directly in front of the basket/goal and 2-point and 3-point positions on the right side of the court at a 90° angle to the basket/goal.

2.3.2. Dependent

The dependent variables for the experiment included subjective and objective measures of presence. Presence was measured subjectively at the end of each trial by using the two-item PQ developed by Draper and Blair (1996). The items in the questionnaire included: (PQ1) "I felt as though I was actually in the synthetic environment as I performed the task" and (PQ2) "The experience involved unity or fusion of self with the synthetic environment". A seven-point rating scale was associated with each item and was used to capture the degree to which a user agreed with the statements (i.e., subjective ratings of presence were made).

Presence was also objectively measured in terms of secondary-monitoring task performance (SD rates) during the trials. Draper et al. (1998) hypothesized that increases in allocation of attentional resources to VR stimuli, as compared to stimuli present in a user's surrounding reality, might lead to increases in telepresence. In this experiment, the secondary tasks randomly presented visual cues irrelevant to the primary task (photoflashes in the stands of the virtual stadium, and strobe light flashes in our real research lab). Telepresence was quantified as the ratio of performance in the photoflash detection versus strobe light

detection; i.e., objective telepresence was equated to VR_{SD}/RW_{SD}. This measure is very similar to the attention allocation measure used by Riley et al. (2004). A high ratio ($VR_{SD} > RW_{SD}$, or $VR_{SD}/RW_{SD} \geqslant 1.0$) was considered indicative of increased attention allocation to the VR and a greater potential for presence experiences.

We also measured subjective workload after each session by using a mental demand rating scale with anchors of "Low" and "High". Mental demand in the task was defined based on the NASA-Task Load Index (TLX) (Hart and Staveland, 1988); i.e., how much thinking, deciding, looking, etc. was required and was the task easy or demanding, simple or complex, and exacting or forgiving. Participants marked an "X" on the scale at the position they felt most accurately represented the demand for the trial. The response measure was the distance from the "Low" anchor to the participant's rating divided by the total length of the scale.

Finally, virtual basketball performance was measured as the number of successful shots made in a 2-min period and as the percentage of successful shots to total shots attempted. The number of baskets/goals during trials was displayed in real-time in the text frame within the view volume so participants could see the results of their performance. The sporting task simulation recorded operator performance automatically.

2.4. Experimental design

A mixed design (between and within variables) was used for this experiment. An equal number of participants (eight) were randomly assigned to groups according to the four levels of the viewpoint condition (a between-subject variable). The auditory cues and visual background of the VR were controlled as within-subject variables. The entire experiment design was replicated; therefore, each participant completed two trials under each of the eight experimental conditions (four sound levels × two visual backgrounds), totaling 16 trials. Each participant experienced half of the test trials at the 2-point distance and half at the 3-point distance. Similarly, each participant completed eight trials with a direct shot to the basket and eight trials shooting baskets from the right side of the court (at a 45° angle from the top of the "key"). The settings of the direct shots and from the right side of the court were balanced across the test trials for two participants. Consequently, the design yielded two replicates for each sound, visual background and distance combination, and four replicates under each sound and visual background combination per pair of participants. The experimental condition setup was the same for all four viewpoints.

2.5. Procedure and training

The procedures for the experiment included: (1) 10 min for anthropometric data collection and familiarization with the test devices; (2) 10 min for administration of an Immersive Tendencies Questionnaire (ITQ) (Witmer and Singer, 1994) to characterize participant susceptibility to presence experiences in advance of the experiment; (3) 20 min for familiarization with the simulation sounds (e.g., ball bouncing, swish, crowd noise, etc.) and visual backgrounds, as well as the telepresence survey (PQ) and mental workload rating scale; (4) 30 min of task training, including learning to control the simulated basketball player and making a shot within the VR by using the six degree-of-freedom mouse, and completion of four 2-min training trials (see detailed description below); (5) 10 min for familiarization with the secondary tasks; (6) 120 min for completion of the 16 test trials with each followed by the PQ, the workload rating, and a 5-min break; and (7) 15 min for completion of a Simulator Sickness Questionnaire (SSQ; Kennedy et al., 1993), before and after training, after the first eight trials, and after all of trials. The first SSQ rating was used as a baseline reading. The SSQ was used to assess participant' well-being and safety, and the data were not analysed for the study. In general, no participants experienced symptoms included in the SSQ that triggered action, based on Kennedy et al. criteria. The experiment lasted approximately 3.5 h for each participant.

Before participants began the experiment, they were provided with a dedicated training session, in which they experienced their randomly assigned viewpoint (according to the experiment design). They heard relevant sounds, as part of the simulation (excluding cheering), and viewed a rendered visual of the basketball stadium (without an audience). Participants made shots on the virtual basket from a 45° angle to the left side of the court from a distance between the 2- and 3-point shot lines. They were permitted four 2-min periods for these shots.

2.6. Specific hypotheses

With respect to the viewpoint condition, the egocentric view was expected to yield the highest presence ratings because it provided participants with the greatest degree of immersion in the environment and performance of the task (Psokta and Davidson, 1993). In regard to performance, the egocentric condition was not expected to produce the highest scores because the user was not provided with the best view of the distance to the backboard, as compared with the exocentric view from the sidelines. The sideline exocentric view allowed participants to see the axis that they were working in (z-axis or depth) and enabled them to better visualize the distance between the player and the goal. The selectable viewpoint was expected to provide high performance and presence ratings because the participant could exploit the advantages of each viewpoint.

It was expected that the addition of either task-relevant or irrelevant auditory stimuli would contribute to the sense of presence in the VR. The task-relevant cues were expected to be more important to presence ratings than task-irrelevant information. This is because the former provides users with the sense that their actions have an

impact on the VR and this is another factor Sheridan (1992) identified as being critical to presence in his model. We also expected that task-irrelevant sound might compromise performance in comparison to the no sound condition and task-relevant sound because of possible distraction from the latter setting. It was also expected that task-relevant sounds would improve performance, as a result of feedback on participant actions in the simulation.

Based on the results of Barfield et al. (1995), presence was expected to be greater under the rich visual background condition because the user had a more realistic view of the environment. It was expected that there would be no difference in performance between the rich and simple visual backgrounds. There were no cues provided in either background that directly aided in the completion of the basketball-shooting task.

Finally, it was expected that interaction effects among the independent variables might be important to participants. That is, the impact of presenting task-irrelevant sounds (e.g., fans booing a "miss") on presence might be moderated by whether a user was provided with clear visuals of the objects that produce such sounds (i.e., a cheering crowd). Therefore, presence ratings under the irrelevant auditory cueing condition were expected to be greater with the rich visual background presenting texture maps of a cheering crowd, etc. If the cues corresponded with each other, or made sense relative to user experiences in reality, we thought this would have a greater positive impact on presence experiences.

3. Data analysis and results

Multi-way analyses of variance (ANOVAs) were applied to the various dependent variables to investigate the influence of viewpoint, sound, visual background and task difficulty (shot distance) on the sense of presence, objective presence, workload and task performance. For the ANOVAs on the presence measures, if shot distance did not prove to be significant in the complete statistical model, a reduced model excluding distance (and pooling variance terms) was used. The full model included four main effects, a subject term, and all two-, three- and four-way interactions defined by full crossings of the settings of the various independent variables. The model accounted for 127 degrees-of-freedom in each data set and an error term was constructed from the complete replicate of the experiment design.

3.1. Presence

The immersive tendencies of all participants were recorded at the outset of the experiment using Witmer and Singer's (1994) ITQ measure. The questionnaire revealed a mean participant score of 60 (out of 126 possible rating points) with a standard deviation of 10.3. This measure is considered later in the results as part of a correlation analysis with presence ratings collected at the end of test trials.

The ANOVAs on the full statistical model (including viewpoint, sound, visual background and task difficulty variables) revealed no significant effect of task difficulty/shot distance on the subjective ratings of presence. Consequently, this variable was dropped from the model, and the reduced model was used to assess the influence of the other VR factors on PQ ratings. ANOVA results on the reduced model revealed significant main effects of immersiveness (viewpoint) ($F(3,63)=4.82$, $p<0.01$), auditory cue type ($F(3,63)=13.44$, $p<0.0001$) and individual differences ($F(4,63)=4.88$, $p<0.01$) on presence ratings (a composite score, including the ratings for PQ1 and PQ2). There was no significant effect of background fidelity or interaction effects (including those involving the subject term as part of the statistical model) on the composite presence measure. Fig. 4 presents the mean presence ratings across the levels of the sound condition and the viewpoint condition for the overall PQ measure. The sound and viewpoint condition means and standard deviations for the presence ratings are included in Table 1. In general, the subjective ratings of presence increased as a greater number of audio cues were presented in the VR and, in particular,

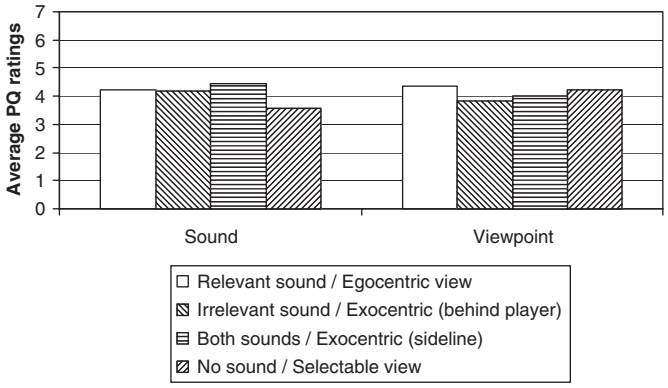

Fig. 4. Average presence ratings across sound and viewpoint conditions.

Table 1
Means and standard deviations (in parentheses) for telepresence and workload ratings for all sound and viewpoint settings

Independent variables	Dependent variables	
	Telepresence	Workload
Sound		
Relevant sound	4.24 (1.12)	53.95 (21.71)
Irrelevant sound	4.17 (1.15)	56.85 (21.34)
Both sounds	4.45 (1.13)	57.27 (21.15)
No sound	3.57 (1.27)	49.41 (23.48)
Viewpoint		
Egocentric view	4.37 (0.90)	55.59 (19.72)
Exocentric (behind player)	3.85 (1.06)	42.79 (22.39)
Exocentric (sideline)	4.00 (1.70)	57.89 (26.04)
Selectable view	4.21 (0.96)	61.20 (14.14)

as task-relevant audio cues were added. With respect to the viewpoint conditions, it can be observed from the graph that the most immersive viewpoint and the viewing option that allowed participants to select among all viewpoints produced higher perceived presence.

Tukey's Honestly Significant Difference tests on the presence ratings were conducted to further investigate the significant sound and viewpoint main effects. The post hoc procedure revealed significantly higher ($p<0.05$) ratings under the sound condition including both relevant and irrelevant cues, the condition presenting only relevant sounds, and the irrelevant sound condition, as compared with the no sound condition. According to Tukey's tests, the egocentric viewpoint and selectable viewpoint produced significantly higher ($p<0.05$) presence ratings than the exocentric view from the sideline, and the exocentric view from behind the player. This made sense as participants also had access to the more immersive egocentric viewpoint under the selectable, or preferred, viewpoint condition.

Contrary to these results on subjective presence, there appeared to be no significant effects of audio cue type, visual background or viewpoint on the objective (attention-based) measure of presence defined in terms of secondary (monitoring) task performance in the VR and RW. The reduced statistical model, excluding the task difficulty factor, also indicated no significant interaction effects on the ratio of VR_{SD}/RW_{SD}.

3.2. Workload

An ANOVA on the full statistical model revealed significant effects of viewpoint ($F(3,127)=16.23$, $p<0.0001$) and audio cue type ($F(3,127)=3.28$, $p<0.05$) on subjective workload captured using the uni-dimensional rating scale. Individual differences within viewpoint condition were also significant ($F(4,127)=4.65$, $p<0.01$) in effect on workload ratings. There were no interaction effects among VR features on workload.

Fig. 5 presents a plot of the mean workload ratings for each viewpoint condition. In general, it can be observed from the graph that the immersive viewpoint and the viewing option that allowed participants to select among all viewpoints produced higher perceived workload in

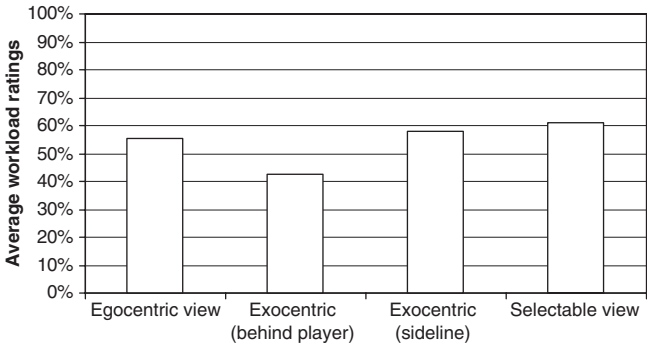

Fig. 5. Average workload ratings across viewpoint conditions.

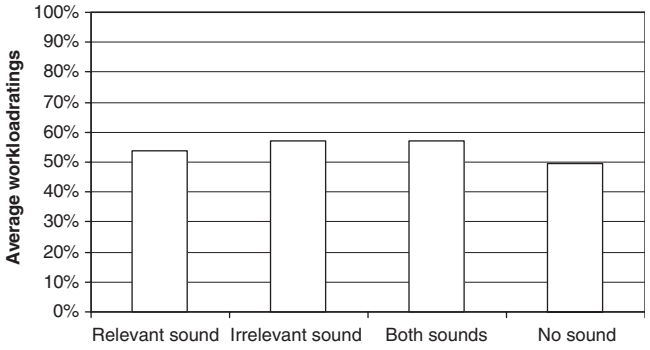

Fig. 6. Average workload ratings across sound conditions.

comparison to the different exocentric viewpoint conditions (the same conditions that generated higher telepresence ratings). Fig. 6 presents a plot of the mean workload ratings for each audio cue type condition. In general, subjective ratings of workload increased as a greater number of audio cues were presented in the VR and as irrelevant cues were added. The means and standard deviations for the workload ratings for all viewpoint and sound conditions are also presented in Table 1.

Tukey's Honestly Significant Difference test was used to further investigate the significant workload effects. Results indicated perceived mental demand to be significantly greater ($p<0.05$) when participants used the selectable (or preferred) viewpoint, exocentric view from the sideline and egocentric viewpoint, as compared with the exocentric view from behind the player. With respect to the audio cue type, significantly lower ($p<0.05$) ratings of workload were recorded under the no cue condition, as compared to the use of relevant, irrelevant, and both relevant and irrelevant sounds in the VR.

3.3. Performance

The results on an ANOVA on the complete statistical model revealed a significant effect due to the task difficulty (shot distance) in terms of total baskets ($F(1,127)=17.62$, $p<0.0001$) and the ratio of total baskets to total shots ($F(1,127)=11.26$, $p<0.001$). Contrary to our expectations, there were no main effects of the other VR system factors, or interaction effects on performance.

The ANOVA indicated that there were a significantly greater number of baskets/goals at the 2-point distance than at the 3-point distance. There was an average 0.93 baskets at the 2-point distance, and 0.54 baskets at the 3-point distance. In general, the total number of baskets was a more sensitive measure of performance in this analysis than the number of baskets per shots attempted.

3.4. Correlation analyses

Simple correlation analyses were also conducted in order to identify any significant relationships among presence, workload, virtual task performance, secondary-task

performance (the hit-to-signal ratio for the camera flash detection task and the hit-to-signal ratio for the strobe light detection task) and immersive tendency scores captured using the ITQ. A Pearson correlation coefficient revealed a significant positive linear association between the overall ITQ score and the subjective ratings of presence ($r = 0.24$, $p < 0.0001$), but not among ITQ scores and the objective measure of presence, which was based on secondary-task performance.

As expected, a Pearson correlation coefficient revealed a significant positive association of subjective presence with workload ratings ($r = 0.39$, $p < 0.0001$). The objective measure of presence was also significantly, but weakly, correlated with perceptions of workload ($r = 0.15$, $p < 0.01$). We considered this to be evidence that the secondary tasks may have operated as objective indicators of workload versus describing attention allocation for achieving presence. Relevant to this, we did find that subjective workload was also significantly, negatively related to the hit-to-signal ratio for the strobe light detection task ($r = -0.21$, $p < 0.0001$).

Pearson correlation coefficients also established a relation of workload and performance in the study. The mental workload ratings were significantly, but weakly, negatively related to performance (total baskets) ($r = -0.12$, $p < 0.05$).

4. Discussion

The results of this study provide evidence that the features of an SE, which include immersiveness (viewpoint) and auditory cue type, significantly influence the sense of subjective presence. These results are consistent with the findings of Slater et al. (1996) and Dinh et al. (1999). However, our findings did not support those of Barfield et al. (1995) indicating that increased visual realism of an SE increases presence. It is possible that there is some minimum level of visual realism necessary for the sense of presence in VR; however, beyond that level, identifying differences among comparatively rich visual backgrounds in terms of presence may require more than rendering versus texturing and presentation of additional objects.

Unfortunately, in this study, the objective attention-based measure of presence did not prove to be sensitive to the experimental manipulations, like subjective ratings of presence. This differs from previous findings, e.g., Riley and Kaber (2001) found a significant correlation between visual attention to VR displays versus elements of the surrounding reality and ratings of telepresence (PQ scores using Witmer and Singer's, 1994 measure). This difference may be attributable to the unique design of the secondary tasks as part of the measure used in this study or differences in the experiment condition manipulations. However, as Draper et al. (1998) have observed, and as Witmer and Singer (1994, 1998) demonstrated, since telepresence is a subjective experience that is significantly influenced by individual susceptibility, it may be that telepresence is most sensitively measured using subjective methods or ratings.

This study indicated level of task difficulty to be an important factor in performance in VR, but there was no corresponding significant effect of task difficulty on presence ratings. This finding also differs from the results of the studies by Riley and Kaber (2001) and Riley et al. (2004). They explored multiple levels of difficulty in simulated teleoperation tasks, e.g., requiring participants to search for and detect land mines in an outdoor VR using a robotic rover. In Riley and Kaber's (2001) study, participants who were exposed to a high difficulty condition reported significantly lower ($p < 0.05$) telepresence (possibly as a result of task overload) than participants in low or moderate groups, which were not significantly different.

The results of the present study also indicated that viewpoint, sound and individual differences were significant factors in subjective workload. We suspected that the observed influence of the viewpoint condition on workload (as well as presence ratings) might also be attributable to the nature of the task, as dictated by the viewpoint (i.e., ease of observing the distance between the player and the goal). Mental workload was also significantly, positively correlated with subjective presence. This is consistent with the findings of the study by Draper and Blair (1996), in which telepresence ratings were significantly correlated with composite workload scores during completion of a pipe-cutting task using a teleoperator. In our study, all VR factor settings that led to greater perceptions of presence also led to increased perception of mental demand. This is supported by the various statistical analyses. Other previous research (Riley and Kaber, 2001) has shown that telepresence may be significantly, negatively correlated with mental workload. However, some of the simulated teleoperation conditions in Riley and Kaber's (2001) study may have led to very high ratings of frustration, as part of the NASA-TLX, because of difficulty participants had in finding mines with the rover. This would have led to high overall workload scores and possible disengagement of participants from the task due to frustration. In general, all these results taken together indicate that some minimum level of workload may be necessary to develop a sense of presence; however, when workload is too high, it may degrade telepresence. Finally, in the present study, the mental workload ratings were significantly, negatively related to performance. It appeared that participants performed worse, although they worked hard (such as thinking and calculating the force to shoot a basket) in the simulation.

This study also revealed a significant correlation between ITQ and the subjective ratings of presence, but not among ITQ scores and the objective measure of presence, based on secondary-task performance. These results were not unexpected as the attention-based objective measure of presence did not appear to be a sensitive indicator of changes in presence due to the manipulations of factors

anticipated to underlie presence experiences (which were revealed by the subjective ratings). Rather, the ratio of SD performance in secondary tasks in VR and reality appeared to be more of an indicator of the primary task load than attention allocation to displays toward achieving presence. The findings on the objective presence measure also indicated that Draper and Blair's (1996) simple PQ may be a quick and valid subjective assessment of presence relative to immersive tendencies identified by Witmer and Singer (1994).

The subjective presence measure was not correlated with virtual task performance. This might have been expected based on the primary ANOVA results of the study demonstrating task difficulty to be important to performance, but sound and viewpoint to dictate presence ratings. This differs from the results of Zhang et al.'s (2005) recent study, in which they found the addition of 3-D auditory or visual feedback to improve virtual assembly task performance. Unfortunately, they did not measure presence in the virtual task. Although the shot distance in our task significantly influenced task performance, it did not appear to influence presence ratings. The study also did not reveal a correlation of objective presence and performance. These results differ from the findings of previous research showing significant correlations of presence and performance (Kaber et al., 2000; Riley and Kaber, 2001). The differences maybe attributed to the specific experimental manipulations or the nature of the task. In this study, we selected independent factors and settings that created the potential for presence and performance to function in different ways in order to, in part, verify presence as a unique construct. It is also possible that since the basketball simulation primarily required psycho-motor performance (distance estimation, shooting the ball) and required more participant interaction skills, presence may not have been as critical to performance as in a telerobot control scenario, like that explored by Kaber et al. (2000), requiring cognitive functions including manipulator path planning for which immersion of one self in an SE interface may be key to performance.

There were no significant interaction effects among all factors (visual background, sound, viewpoint and task difficulty) manipulated in this study in terms of performance, presence and workload. Although counter to expectation, and our hypothesis that a cue conflict might occur as a result of auditory cue and visual background manipulations (i.e., a cheering crown with no visual of an audience in the virtual stadium), this result is consistent with the findings by Dinh et al. (1999). They investigated the effects of tactile, olfactory, audio and visual sensory cues on a participant's sense of presence in a virtual environment that consisted of a corporate office suite including a reception area, hallway, bathroom, small office, copier room, large office and balcony. Their analysis revealed no significant interaction effects on subjective measure of presence. In the present study, information from the different sensory channels (e.g., auditory and visual) did not interact to improve or decrease overall performance and presence in the VR, nor did the combination of cues influence the subjective perception of workload. These results do not support the intuitive hypothesis on cross-modal interaction in VR by Biocca et al. (2001) that multiple sensory cueing of events in an SE will lead to increased presence experiences. However, the current study did not explore haptic cues, like Biocca et al. (2001). It is possible that there is some interaction between the visual and haptic modalities that may be critical to presence in VR tasks for which motor control is a significant part. It may be more likely, as Stanney (2003) contends, that the interaction (or general importance) of multimodal cues in VR design depends on the demands of the task and designers may need to identify those cues that best support performance.

5. Conclusions

The findings presented above provide some basic insight into the potential presence implications of the design of VR or SEs for use in teleoperation tasks/systems; however, the results may be limited in terms of generalizability to specific real-world teleoperation scenarios. There is a need for elaborate investigations of the role of multimodal SE design factors in telepresence, performance and workload using tasks that are more representative of realistic SE applications. Contrary to our initial projection regarding the use of the basketball simulation in this study, it is possible that in a more complex VR application, involving rich sounds and visual information, presence experiences and fluctuations due to changes in system interface features may be even greater. As an example, in a teleoperated surgical environment implemented by Freysinger et al. (2002), highly detailed audio and visual information was provided to users to promote a higher sense of telepresence and to promote operator performance. The absence of this information in the integrated SE could be detrimental to a surgeon's perception of involvement in an actual surgical operation and ultimately the accuracy of performance. Based on the present research, the use of auditory cues in SEs for telerobot control may serve to enhance user telepresence and performance. Furthermore, the viewpoint of a user may be critical to telepresence in the teleoperation environment.

One concern that can be raised, based on the findings presented here with respect to design for telepresence in complex teleoperation applications, is that the manner in which SE features combine or interact together (e.g., auditory cues, immersive viewpoint, visual fidelity) may not be additive in terms of a resulting affect on telepresence and synthetic task performance. Based on our results, the lack of interaction effects of SE factors on subjective presence measures and VR task performance suggests that design using multiple sensory channels could be limited to consideration of main effects of each channel on

performance, presence, etc., as part of the interface development process. This needs to be further evaluated through observation of VR user experiences in higher fidelity simulations involving realistic SE applications.

Caution may be necessary in manipulating multiple sensory channels and features in SE-based interface design for complex systems to facilitate presence experiences, as there may be a mental workload "side effect". In the present study, the pattern of ANOVA results on, and correlations among, subjective mental demand ratings and telepresence in the virtual sporting task supported this contention. It is important to note, however, that mental workload was generally low across all experimental conditions (on a scale from 0% to 100%, all condition means were less than 60%). It is possible that users may be even more sensitive to SE feature manipulations and combinations of various viewpoint, auditory cue and visual background settings under more demanding teleoperation task circumstances.

Finally, the lack of correlation of presence experiences and performance in the present study should also be of concern is designing actual teleoperation systems. Since the virtual sporting we used primarily required psycho-motor behaviors of participants, it may not have been as sensitive of a paradigm for revealing a relation among the constructs, as compared to a teleoperation simulation requiring complex cognitive processes for manipulator path planning and tracking. These types of task performance requirements are critically dependent upon user achievement of situation awareness on the state of a robotic system and remote task environment that, in-turn, may be dependent upon users achieving self-fusion or telepresence with an SE-based interface.

Acknowledgments

This research was supported by a grant from the National Science Foundation (NSF) entitled, "CAREER: Telepresence in Teleoperations" (Grant No. IIS-0196342). The opinions expressed in this paper are those of the authors and do not necessarily reflect the views of the NSF. Laura Hughes, Mohamed Sheik-Nainar, Melanie Wright and John Draper made input in the early design stage of the study. The authors would also like to thank Mohamed Sheik-Nainar, who contributed to the development of the task simulation. We also thank Srikanth Jayanthi who assisted with the experimental data collection. Finally, we are grateful to the three anonymous reviewers for their comments and suggestions.

References

Barfield, W., Hendrix, C., Bjorneseth, O., 1995. Spatial performance with perspective displays as a function of computer graphics eyepoint elevation and geometric field of view. Applied Ergonomics 26, 307–314.

Biocca, F., Kim, J., Choi, Y., 2001. Visual touch in virtual environments: an exploratory study of presence, multimodal interface, and cross-modal sensory illusions. Presence 10 (3), 247–265.

Dinh, J.Q., Walker, N., Hodges, L.F., Song, C., Kobayashi, A., 1999. Evaluating the importance of multi-sensory input on memory and the sense of presence in virtual environments. In: Proceedings of the Virtual Reality Annual International Symposium. IEEE Computer Society, Houston, TX, pp. 222–228.

Draper, J.V., Blair, L.M., 1996. Workload, flow, and telepresence during teleoperation. In: Proceedings of the International Conference on Robotics and Automation. Minneapolis, MN, pp. 1030–1035.

Draper, J.V., Kaber, D.B., Usher, J.M., 1998. Telepresence. Human Factors 40 (3), 354–375.

Draper, J.V., Kaber, D.B., Usher, J.M., 1999. Speculations on the value of presence. Cyber Psychology and Behavior 2 (4), 349–361.

Freysinger, W., Truppe, M.J., Gunkel, A.R., Thumfart, W.F., 2002. Stereotactic telepresence in ear, nose, and throat surgery. HNO 50 (5), 424–432.

Hart, S.G., Staveland, L.E., 1988. Development of NASA-TLX (Task Load Index): results of empirical and theoretical research. In: Hancock, P.A., Meshkati, N. (Eds.), Human Mental Workload. North-Holland Elsevier Science, Amsterdam, pp. 139–183.

Kaber, D.B., Riley, J.M., 2000. Task difficulty and user motivation effects on performance, and telepresence in a teleoperation task. In: Proceedings of the Third International Conference on Engineering Psychology and Cognitive Ergonomics. Ashgate, Aldershot, UK.

Kaber, D.B., Riley, J.M., Zhou, R., Draper, J.V., 2000. Effects of visual interface design, and control mode and latency on performance, telepresence and workload in a teleoperation task. In: Proceedings of the XIVth Triennial Congress of the International Ergonomics Association/44th Annual Meeting of the Human Factors and Ergonomics Society. San Diego, CA, August 2000.

Kennedy, R.S., Lane, N.E., Berbaum, K.S., Lilienthal, M.G., 1993. Simulator Sickness Questionnaire: an enhanced method for quantifying simulator sickness. International Journal of Aviation Psychology 3 (3), 203–220.

Nash, E.B., Edwards, G.W., Thompson, J.A., Barfield, W., 2000. A review of presence and performance in virtual environment. International Journal of Human–Computer Interface 12 (1), 1–41.

Nichols, S., Haldane, C., Wilson, R., 2000. Measurement of presence and its consequences in virtual environment. International Journal of Human–Computer Studies 52, 471–491.

Psokta, J., Davidson, S., 1993. Cognitive Factors Associate with Immersion in Virtual Environments. US Army Research Institute, Alexandria, VA.

Riley, J.M., 2001. Investigating the utility of dynamic measures of situation awareness for quantifying telepresence experiences in a teleoperation task. Unpublished Doctoral Dissertation, Mississippi State University, Starkville, MS.

Riley, J.M., Kaber, D.B., 2001. Utility of situation awareness and attention for describing telepresence experiences in a virtual telepresence task. In: Das, B., Karwowski, W. (Eds.), Proceedings of the 2001 International Conference on Computer-Aided Ergonomics and Safety (CD-ROM). Maui, HI, July 28–August 2, 2001.

Riley, J.M., Kaber, D.B., Draper, J.V., 2004. Using measures of situation awareness and attention allocation to quantify telepresence experiences in teleoperation. Human Factors & Ergonomics in Manufacturing 14 (1), 51–68.

Sadowski, W., Stanney, K.M., 2002. Presence in virtual environments. In: Stanney, K.M. (Ed.), Handbook of Virtual Environments. LEA, Mahwah, NJ, pp. 791–806.

Sas, C., O'Hare, M.P.G., 2003. Presence equation: an investigation into cognitive factors underlying presence. Presence: Teleoperators and Virtual Environments 12 (5), 523–537.

Sheik-Nainar, M.A., Kaber, D.B., Chow, M.-Y., 2005. Control gain adaptation in virtual reality mediated human–telerobot interaction. Human Factors & Ergonomics in Manufacturing 15 (3), 249–332.

Sheridan, T.B., 1992. Musings on telepresence and virtual presence. Presence: Teleoperators and Virtual Environments 1 (1), 120–126.

Slater, M., Steed, A., 2000. A virtual presence counter. Presence: Teleoperators and Virtual Environments 9 (5), 413–434.

Slater, M., Wilbur, S., 1997. A framework for immersive visual environment: speculations on the role of presence in virtual environments. Presence: Teleoperators and Virtual Environments 6 (6), 603–617.

Slater, M., Linakis, V., Usoh, M., Kooper, R., 1996. Immersion, presence and performance in virtual environments: an experiment with tri-dimensional chess. In: Mark, G. (Ed.), ACM Virtual Reality Software and Technology (VRST), pp. 163–172.

Stanney, K.M., 2003. Virtual environments. In: Jacko, J., Sears, A. (Eds.), The Human–Computer Interaction Handbook. Erlbaum, Mahwah, NJ, pp. 622–632 (Chapter 31).

Stanney, K.M., Mourant, R.R., Kennedy, R.S., 1998. Human factors issues in virtual environments. Presence: Teleoperators and Virtual Environments 7 (4), 327–351.

Welch, R.B., 1999. How can we determine if the sense of presence affects task performance? Presence 8 (5), 574–577.

Witmer, B.G., Singer, M.J., 1994. Measuring immersion in virtual environments (Technical Report 1014). US Army Research Institute for the Behavioral and Social Science, Alexandria, VA.

Witmer, B.G., Singer, M.J., 1998. Measuring presence in virtual environments: a Presence Questionnaire. Presence 7 (3), 225–240.

Zhang, Y., Sotudeh, R., Fernando, T., 2005. The use of visual and auditory feedback for assembly task performance in a virtual environment. In: Proceedings of ACM, 1-59593-203-6/05/0005, pp. 59–66.